

# APPENDIX A

PROSITE documentation PDOC00124

[ExPASy Home page](#) [Site Map](#) [Search ExPASy](#) [Contact us](#) [PROSITE](#)

Search  for

**prosite** [Home](#) [ScanProsite](#) [ProRule](#) [Documents](#) [Downloads](#) [Links](#)

PROSITE documentation PDOC00124 

## Serine proteases, trypsin family, signatures and profile

### Description:

The catalytic activity of the serine proteases from the trypsin family is provided by a charge relay system involving an aspartic acid residue hydrogen-bonded to a histidine, which itself is hydrogen-bonded to a serine. The sequences in the vicinity of the active site serine and histidine residues are well conserved in this family of proteases [1]. A partial list of proteases known to belong to the trypsin family is shown below.

- Acrosin.
- Blood coagulation factors VII, IX, X, XI and XII, thrombin, plasminogen, and protein C.
- Cathepsin G.
- Chymotrypsins.
- Complement components C1r, C1s, C2, and complement factors B, D and I.
- Complement-activating component of RA-reactive factor.
- Cytotoxic cell proteases (granzymes A to H).
- Duodenase I.
- Elastases 1, 2, 3A, 3B (protease E), leukocyte (medullasin).
- Enterokinase (EC 3.4.21.9) (enteropeptidase).
- Hepatocyte growth factor activator.
- Hepsin.
- Glandular (tissue) kallikreins (including EGF-binding protein types A, B, and C, NGF- $\gamma$  chain,  $\gamma$ -renin, prostate specific antigen (PSA) and tonin).
- Plasma kallikrein.
- Mast cell proteases (MCP) 1 (chymase) to 8.
- Myeloblastin (proteinase 3) (Wegener's autoantigen).
- Plasminogen activators (urokinase-type, and tissue-type).
- Trypsins I, II, III, and IV.
- Tryptases.
- Snake venom proteases such as anrod, batroxobin, cerastobin, flavoxobin, and protein C activator.
- Collagenase from common cattle grub and collagenolytic protease from Atlantic sand fiddler crab.
- Apolipoprotein(a).
- Blood fluke cercarial protease.
- Drosophila trypsin like proteases:  $\alpha$ , easter, snake-locus.
- Drosophila protease stubble (gene sb).
- Major mite fecal allergen Der p III.

All the above proteins belong to family S1 in the classification of peptidases [2,E1] and originate from eukaryotic species. It should be noted that bacterial proteases that belong to family S2A are similar enough in the regions of the active site residues that they can be picked up by the same patterns. These proteases are listed below.

- *Achromobacter lyticus* protease I.
- *Lysobacter*  $\alpha$ -lytic protease.
- Streptogrisin A and B (*Streptomyces* proteases A and B).
- *Streptomyces griseus* glutamyl endopeptidase II.
- *Streptomyces fradiae* proteases 1 and 2.

We also developed a profile specific for the S1 family that spans the complete domain. In addition to proteases from the S1 family, this profile also detects proteins that have lost active site residues and which are therefore no longer catalytically active. Examples of such proteins are haptoglobin and protein Z.



## Note:

If a protein includes both the serine and the histidine active site signatures, the probability of it being a trypsin family serine protease is 100%.



## Last update:

May 2002 / Text revised.

## Technical section:

PROSITE methods (with tools and information) covered by this documentation:

## TRYPSIN\_DOM: PS50240: Serine proteases\_trypsin\_domain\_profile (MATRIX)

*Sequences known to belong to this class detected by the profile:* ALL

*Other sequence(s) detected in Swiss-Prot:* NONE

- Domain architecture view of Swiss-Prot proteins matching PS50240

- Retrieve an alignment of Swiss-Prot true positive hits:  
Clustal format, color, condensed view / Clustal format, color / Clustal format, plain text / Fasta format
- Taxonomic tree view of all Swiss-Prot/TrEMBL entries matching PS50240
- Retrieve a list of all Swiss-Prot/TrEMBL entries matching PS50240
- Scan Swiss-Prot/TrEMBL entries against PS50240
- view ligand binding statistics

*Matching PDB structures:* 1A0H 1A0J 1A0L 1A2C ... [ALL]



## TRYPSIN\_HIS: PS00134: Serine proteases\_trypsin family\_histidine active site (PATTERN)

*Consensus pattern:*

[LIVM]-[ST]-A-[STAG]-H-C

H is the active site residue

*Sequences known to belong to this class detected by the pattern:* ALL, except for complement components C1r and C1s, pig plasminogen, bovine protein C, rodent urokinase, anrod, gyroxin and two insect trypsins

*Other sequence(s) detected in Swiss-Prot:*

18

- Retrieve an alignment of Swiss-Prot true positive hits:  
Clustal format, color, condensed view / Clustal format, color / Clustal format, plain text / Fasta format
- Taxonomic tree view of all Swiss-Prot/TrEMBL entries matching PS00134
- Retrieve a list of all Swiss-Prot/TrEMBL entries matching PS00134
- Scan Swiss-Prot/TrEMBL entries against PS00134
- view ligand binding statistics

*Matching PDB structures:* 1A0H 1A0J 1A0L 1A2C ... [ALL]



## TRYPSIN\_SER: PS00135: Serine proteases\_trypsin family\_serine active site (PATTERN)



<b>Consensus pattern:</b>	[DNSTAGCT]-[GSTAPIMVQH]-x(2)-G-[DE]-S-G-[GS]-[SAPHV]- [LIVMFYWH]-[LIVMFYSTANQH]
<b>Sequences known to belong to this class detected by the pattern:</b>	ALL, except for 18 different proteases which have lost the first conserved glycine.
<b>Other sequence(s) detected in Swiss-Prot:</b>	8
<ul style="list-style-type: none"> <li>• Retrieve an alignment of Swiss-Prot true positive hits: Clustal format, color, condensed view / Clustal format, color / Clustal format, plain text / Fasta format</li> <li>• Taxonomic tree view of all Swiss-Prot/TrEMBL entries matching PS00135</li> <li>• Retrieve a list of all Swiss-Prot/TrEMBL entries matching PS00135</li> <li>• Scan Swiss-Prot/TrEMBL entries against PS00135</li> <li>• view ligand binding statistics</li> </ul>	
<b>Matching PDB structures:</b> 1A0H 1A0J 1A0L 1A2C ... [ALL]	



## References:

<b>Authors</b>	Brenner S.
<b>Title</b>	<i>The molecular evolution of genes and proteins: a tale of two serines.</i>
<b>1 Source</b>	Nature 334:528-530(1988).
<b>PubMed ID</b>	3136396
<b>DOI</b>	10.1038/334528a0

<b>Authors</b>	Rawlings N.D., Barrett A.J.
<b>Title</b>	<i>Families of serine peptidases.</i>
<b>2 Source</b>	Methods Enzymol. 244:19-61(1994).
<b>PubMed ID</b>	7845208

E1    **Source**    <http://www.expasy.org/cgi-bin/lists?peptidas.txt>

## Copyright:

PROSITE is copyright. It is produced by the Swiss Institute of Bioinformatics (SIB). There are no restrictions on its use by non-profit institutions as long as its content is in no way modified. Usage by and for commercial entities requires a license agreement. For information about the licensing scheme send an email to [license@ib-sib.ch](mailto:license@ib-sib.ch) or see: [http://www.expasy.org/prosite/prosite\\_license.htm](http://www.expasy.org/prosite/prosite_license.htm).

## Miscellaneous:

[View entry in original PROSITE document format](#)

[View entry in raw text format \(no links\)](#)